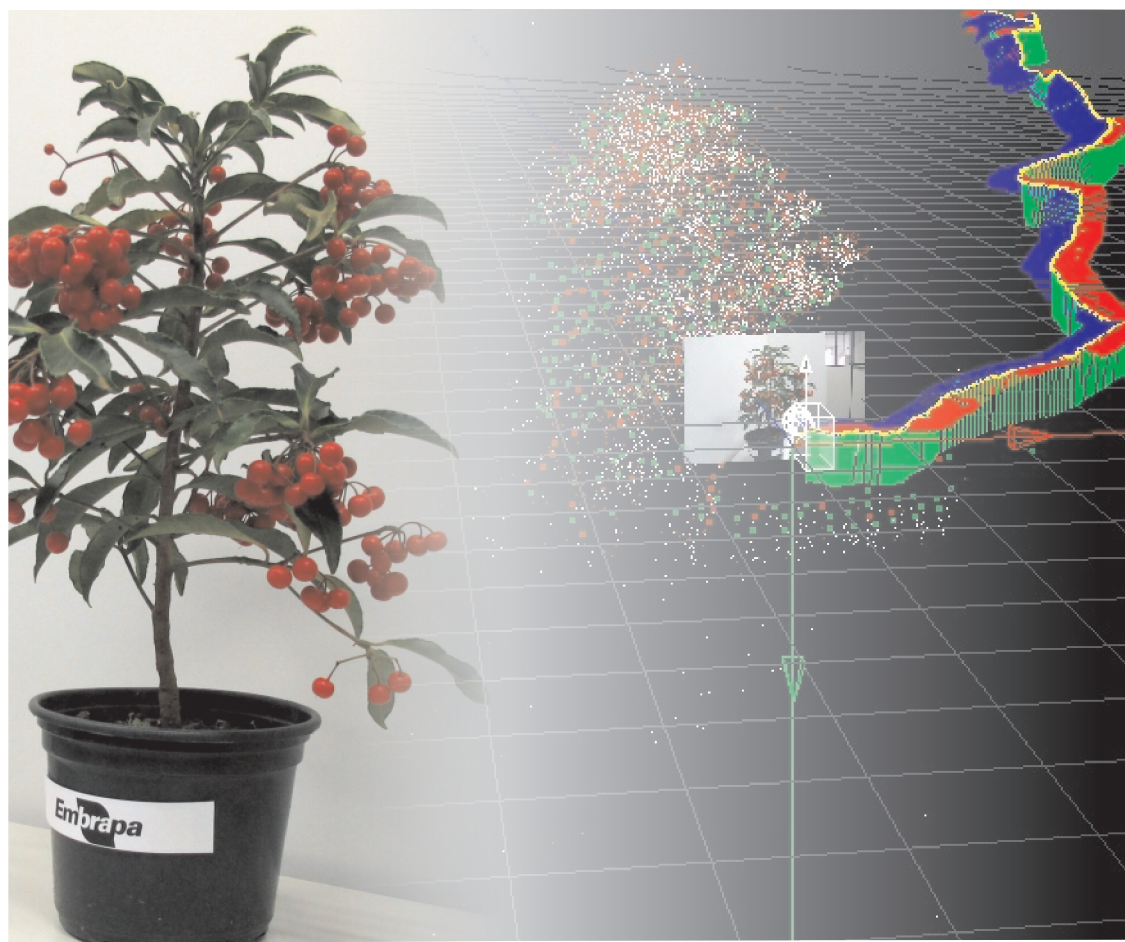


## Digitalização 3D de plantas por imagem para análise de arquitetura e fenótipo





*Empresa Brasileira de Pesquisa Agropecuária  
Embrapa Informática Agropecuária  
Ministério da Agricultura, Pecuária e Abastecimento*

# ***Documentos 119***

## **Digitalização 3D de plantas por imagem para análise de arquitetura e fenótipo**

*Thiago Teixeira Santos*

Embrapa Informática Agropecuária  
Campinas, SP  
2011

## **Embrapa Informática Agropecuária**

Av. André Tosello, 209 - Barão Geraldo  
Caixa Postal 6041 - 13083-886 - Campinas, SP  
Fone: (19) 3211-5700 - Fax: (19) 3211-5754  
www.cnptia.embrapa.br  
sac@cnptia.embrapa.br

## **Comitê de Publicações**

Presidente: *Silvia Maria Fonseca Silveira Massruhá*

Membros: *Poliana Fernanda Giachetto, Roberto Hiroshi Higa, Stanley Robson de Medeiros Oliveira, Maria Goretti Gurgel Praxedes, Adriana Farah Gonzalez, Neide Makiko Furukawa*

Membros suplentes: *Alexandre de Castro, Fernando Attique Máximo, Paula Regina Kuser Falcão*

Supervisor editorial: *Stanley Robson de Medeiros Oliveira, Neide Makiko Furukawa*

Revisor de texto: *Adriana Farah Gonzalez*

Normalização bibliográfica: *Victor Paulo Marques Simão, Maria Goretti Gurgel Praxedes*

Editoração eletrônica/Capa: *Neide Makiko Furukawa*

Secretária: *Carla Cristiane Osawa*

## **1ª edição on-line 2011**

### **Todos os direitos reservados.**

A reprodução não autorizada desta publicação, no todo ou em parte, constitui violação dos direitos autorais (Lei no 9.610).

### **Dados Internacionais de Catalogação na Publicação (CIP) Embrapa Informática Agropecuária**

---

Santos, Thiago Teixeira.

Digitalização 3D de plantas por imagem para análise de arquitetura e fenótipo / Thiago Teixeira Santos.- Campinas : Embrapa Informática Agropecuária, 2011.

28 p. : il. - (Documentos / Embrapa Informática Agropecuária, ISSN 1677-9274 ; 119).

1. Imagem digital. 2. Computação gráfica. 3. Digitalização de planta. 4. Modelo 3D. 5. Arquitetura de planta. I. Embrapa Informática Agropecuária. II. Título. III. Série.

CDD 006.6 (21 ed.)

---

© Embrapa 2011

# Autor

## **Thiago Teixeira Santos**

Doutor em Ciências da Computação

Pesquisador da Embrapa Informática Agropecuária

Av. André Tosello, 209 - Barão Geraldo

Caixa Postal 6041 - 13083-886 - Campinas, SP

Telefone: 19 3211-5823

e-mail: [thiago@cnptia.embrapa.br](mailto:thiago@cnptia.embrapa.br)



# **Apresentação**

A digitalização de um objeto pode ser entendida como a construção de um modelo composto por um conjunto discreto de amostras, permitindo o armazenamento e o processamento desse objeto por um computador. O presente trabalho trata da digitalização das características macroscópicas de plantas, sua estrutura e geometria.

Diversas análises científicas, entre elas os estudos de correlação fenótipo-genótipo em larga escala, demandam que um grande número de espécimes vegetais seja digitalizado rapidamente. Trata-se portanto de um desafio, já que as técnicas atuais de digitalização de plantas envolvem procedimentos manuais lentos e trabalhosos.

Este trabalho apresenta algumas alternativas para digitalização de plantas, encontradas na literatura recente, que empregam diferentes técnicas de imageamento para obter, de forma automática ou semiautomática, modelos tridimensionais de plantas. Também são apresentadas técnicas de construção de modelos 3D utilizadas em outros domínios que podem ser a base para novos avanços em digitalização de plantas.

Sistemas de digitalização de plantas por imagem, empregando modernas técnicas de reconstrução por visão computacional, podem se tornar uma alternativa eficiente, barata, não-invasiva e não-destrutiva para a produção de um grande volume de dados para análise de fenótipos, modelagem e simulação do desenvolvimento de cultivares. Este trabalho busca apontar alternativas para a construção de tais sistemas.

***Kleber Xavier Sampaio de Souza***

Chefe-geral

Embrapa Informática Agropecuária





# Sumário

<b>Introdução .....</b>	<b>9</b>
<b>Digitalização de plantas .....</b>	<b>11</b>
Varredura a laser .....	12
Digitalização de plantas por visão estéreo .....	14
Noções básicas de visão estéreo .....	15
Reconstrução estéreo de plantas .....	16
<b>Avanços recentes em reconstrução 3D de objetos .....</b>	<b>18</b>
<b>Resultados preliminares utilizando visão estéreo .....</b>	<b>21</b>
<b>Conclusões .....</b>	<b>23</b>
<b>Referências .....</b>	<b>24</b>



# Digitalização 3D de plantas por imagem para análise de arquitetura e fenótipo

---

*Thiago Teixeira Santos*

## Introdução

Digitalização tridimensional de plantas se refere à construção de modelos 3D que permitam medições, análises e simulações dependentes de características geométricas tridimensionais dos vegetais. Atualmente, duas áreas se mostram particularmente interessadas nesse tipo de informação: a) a modelagem funcional-estrutural de plantas; b) a fenotipagem em larga escala de plantas.

A modelagem funcional-estrutural de plantas tem como objetivo reproduzir o padrão de crescimento e diferenciação que determinam a arquitetura da planta. A análise da arquitetura da planta visa compreender como a espacialização de seus processos fisiológicos se relacionam à sua morfogênese (GODIN et al., 2005). Segundo Turnbull (2005), a manipulação dos processos que determinam a arquitetura dos cultivares tem sido uma das principais responsáveis pelo aumento de produtividade na agricultura ao longo da história, perdendo em importância apenas para a descoberta das necessidades nutricionais em plantas. Na modelagem computacional de plantas, representações tridimensionais da arquitetura podem ser utilizadas em diversas simulações como, por exemplo, cálculo de interceptação de radiação (GODIN et al., 1999) ou relações fonte/dreno (LOPEZ et al., 2008).

Plataformas para fenotipagem em larga escala têm surgido nos últimos anos com o objetivo de produzir dados de fenótipo em grande quantidade. Exemplos dessas plataformas incluem PHENOPSIS (GRANIER et al., 2006), empregada em estudos sobre *Arabidopsis thaliana* realizados pelo Instituto Nacional de Pesquisa em Agricultura (INRA) e TraitMill™, plataforma utilizada na avaliação de transgênicos de arroz (*Oryza sativa*) (REUZEAU et al., 2006). Tais plataformas devem empregar técnicas não-destrutivas de aquisição de dados pois precisam gerar medições em diversos momentos ao longo do desenvolvimento da planta.

Na modelagem funcional-estrutural de plantas, a digitalização de espécimes tem sido realizada nos últimos anos com o uso de técnicas de contato que exigem que um operador mova um sensor sobre a superfície do vegetal (COSTES et al., 2003; GODIN et al., 1999). O processo é tedioso, trabalhoso e requer longos períodos de tempo. Godin et al. (1999) reportam que 24 dias de trabalho foram necessários para que um operador, utilizando um rastreador magnético (Polhemus Inc., Cochester, VT, EUA), pudesse digitalizar oito espécimes para seus experimentos em arquitetura de macieiras. Rakocevic et al. (2000) levaram até 7 horas para digitalizar trevos (*Trifolium repens* L.) sobre uma pequena área de 100 cm<sup>2</sup> utilizando o mesmo equipamento.

Se tal aquisição manual de dados constitui uma dificuldade à modelagem funcional-estrutural, no caso da fenotipagem de alto desempenho, trata-se de um entrave proibitivo, devido à necessidade de processar sistematicamente diversos indivíduos de modo uniforme (KAMINUMA et al., 2004). Nas plataformas de fenotipagem, técnicas de processamento de imagens têm sido empregadas para medição e análise (HARTMANN et al., 2011; KAMINUMA et al., 2004; REUZEAU et al., 2006), usufruindo de ambiente controlado e automatizado para imageamento. Porém, os sistemas utilizados nessas plataformas são capazes de realizar apenas medições simples como altura e largura, sem obter qualquer outro dado da estrutura tridimensional da planta, limitando-se a modelos bidimensionais simples do vegetal (HARTMANN et al., 2011). Uma exceção é o sistema de reconstrução 3D por varredura a laser proposto por Kaminuma et al. (2004), capaz de obter superfícies tridimensionais para folhas e pedúnculos em *Arabidopsis*.

Técnicas de digitalização 3D baseadas em varredura a laser ou em visão estéreo têm surgido como alternativas práticas, não-intrusivas e não-destrutivas para a construção de modelos com os quais medições na planta

podem ser feitas automaticamente e de modo consistente. No caso de sistemas baseados em visão estéreo, soma-se ainda a vantagem do baixo custo, dada a possibilidade de usar câmeras digitais comuns no mercado.

O presente trabalho visa apresentar o estado da arte em digitalização por imagem de plantas, suas limitações atuais e possíveis desenvolvimentos que podem ser derivados a partir de avanços recentes no campo da visão computacional. A Seção 2 apresenta trabalhos na literatura que abordaram o problema da digitalização de plantas utilizando técnicas de contato, varredura a laser e visão estéreo. A Seção 3 apresenta alguns artigos que representam o estado da arte em reconstrução tridimensional de objetos, promissores como base de novas metodologias para digitalização de plantas. Testes empregando uma dessas técnicas (estruturação por movimento), foram realizados e um resultado preliminar é apresentado na Seção 4. Finalmente, a Seção 5 conclui o texto com algumas considerações sobre possíveis avanços na área.

## Digitalização de plantas

Os métodos para digitalização de plantas podem ser agrupados em duas classes. Os métodos de contato exigem que algum tipo de “ponteiro” toque a superfície da planta de modo que o modelo tridimensional seja obtido pelo registro de sua posição. Já os métodos “sem contato” empregam sensoriamento por dispositivos como *laser scanners* ou câmeras e constroem o modelo da planta processando os dados obtidos por tais sensores.

O trabalho de Lang (1973) é reconhecido como o primeiro a utilizar um aparato especial para produzir uma representação discreta de plantas (GODIN et al., 2005). Tratava-se de um método de contato que empregava um braço mecânico com potenciômetros capazes de registrar as rotações de suas articulações. Uma desvantagem de tal equipamento é que seu posicionamento pode por si próprio alterar a estrutura do dossel dos espécimes sendo modelados. Outro problema comum é que diversos pontos da estrutura podem ser inacessíveis pelo ponteiro, devido a limitações físicas.

Outro método de contato é o digitalizador sônico utilizado por Sinoquet et al. (1991) em experimentos com cultivares de milho. Esse método empre-

gava três captadores de ultrassom. O ponteiro era um emissor de ultrassom e sua posição era determinada computando-se as distâncias a cada captador, calculadas por intervalos de tempo entre emissão e captação do som em cada sensor, assumindo velocidade constante do som pelo ar. O método era sensível ao vento e mesmo a própria estrutura da vegetação poderia alterar a propagação do som, prejudicando os resultados.

Digitalizadores magnéticos são o método de contato utilizado mais intensamente nos últimos anos (GODIN et al., 1999; RAKOCEVIC et al., 2000). Esses dispositivos produzem um campo magnético que induz correntes elétricas em bobinas existentes dentro do ponteiro, permitindo determinar sua posição e sua orientação. Acurácia de poucos milímetros pode ser obtida (RAKOCEVIC et al., 2000) e o campo magnético não é afetado pelas estruturas vegetais, embora objetos metálicos próximos possam prejudicar as medições.

As técnicas de contato exigem que um operador mova o ponteiro do equipamento utilizado, amostrando a superfície da planta. Embora tais técnicas apresentem a vantagem de permitir ao especialista inserir anotações durante a medição, o processo é tedioso e de longa duração (GODIN et al., 1999), de modo que poucos espécimes são digitalizados a cada experimento. Essa é a principal motivação para o uso de técnicas sem contato baseadas em outras formas de sensoriamento, descritas a seguir.

## Varredura a laser

Modelos tridimensionais podem ser obtidos com o uso de *scanners Light Detection And Ranging* (LiDAR). Tais dispositivos são capazes de medir a distância percorrida por um feixe laser até atingir a superfície de um objeto. A superfície pode ser determinada realizando-se uma varredura com o feixe.

Kaminuma et al. (2004) utilizaram um *scanner laser*<sup>1</sup> para construir modelos tridimensionais para *Arabidopsis thaliana*. Os modelos consistiam na superfície 3D de folhas e pecíolos, representadas como malhas poligonais.

---

<sup>1</sup> O equipamento utilizado foi um scanner VOXELAN HEW-50HS (Hamano Engineering Co., Ltd., Kawasaki, Japão).

As malhas foram utilizadas para determinar quantitativamente dois atributos morfológicos, direção da lâmina foliar e epinastia foliar, de modo a caracterizar o fenótipo de dois ecótipos diferentes de *Arabidopsis*: *Columbia* e *Tsu-0*.

Os autores obtiveram uma boa amostragem das superfícies devido a particularidades na montagem do experimento: a distância e o tamanho da amostra permitiram uma resolução de 0.045 mm por pixel, produzindo assim uma nuvem densa de pontos em 3D. Contudo, montagens diferentes com amostras de dimensões maiores como árvores, arbustos ou mesmo espécimes de *Arabidopsis* em estágios mais avançados de desenvolvimento podem produzir um conjunto esparsos de pontos.

Livny et al. (2010) construíram modelos tridimensionais de árvores a partir de nuvens esparsas de pontos obtidos por um *scanner* LiDAR<sup>2</sup> montado no topo de um automóvel. Os pontos gerados correspondiam a árvores próximas das vias por onde o veículo trafegou. Para produzir um “esqueleto” correspondente à estrutura de cada árvore, o sistema levava em conta critérios como comprimento, espessura, suavidade e densidade dos ramos. Tal esqueleto é um modelo simplificado da estrutura ramificada do vegetal e sua posição no espaço, representado computacionalmente como um grafo dirigido acíclico. Uma primeira aproximação do esqueleto era produzida com o uso do algoritmo de Dijkstra<sup>3</sup>, tomando a nuvem de pontos como vértices do grafo. Utilizando otimização global baseada em mínimos quadrados, o algoritmo de Livny et al. (2010) modifica a posição dos vértices de modo a suavizar a orientação das arestas no grafo. A otimização global é mais robusta a ruído e à não-uniformidade na densidade dos pontos. A superfície da árvore é determinada por cilindros generalizados centrados nas arestas do grafo e cujo raio é determinado por alometria (XU et al., 2007). Os resultados são visualmente convincentes, a aparência do modelo é próxima ao espécime imageado, o que faz o método apropriado para aplicações em computação gráfica. Contudo, a fidelidade do modelo não é, a princípio, apropriada para a realização de medições como as necessárias em estudo de fenotipagem. A metodologia não é capaz de modelar apropriadamente espécimes apresentando grande densidade de

---

<sup>2</sup> O *scanner* utilizado foi um Lynx Mobile Mapper (Optech Inc., Toronto, Canadá).

<sup>3</sup> Detalhes sobre grafos dirigidos acíclicos e o algoritmo de Dijkstra podem ser encontrados em Cormen et al. (2009).

folhas, não sendo capaz de operar em situações nas quais a estrutura de ramos não puder ser satisfatoriamente coberta pelo processo de varredura a laser.

De modo similar a Livny et al. (2010), Preuksakarn et al. (2010) desenvolveram um sistema para construção de modelos tridimensionais para árvores sem presença de folhas. A nuvem de pontos tridimensionais foi gerada a partir de varredura a laser<sup>4</sup> e modificada por um algoritmo de contração que move os pontos da superfície da planta em direção ao eixo central do segmento (GIANNITRAPANI; MURINO, 1999). A nuvem contraída é utilizada como entrada de um algoritmo de colonização de espaço (space colonization algorithm - SCA) proposto por Runions et al. (2007). Esse algoritmo iterativamente expande uma estrutura 3D ramificada (esqueleto) utilizando os pontos da nuvem como “atratores”. O esqueleto se expande até não existirem mais atratores, que são removidos ao serem cobertos pela estrutura. Finalmente, cilindros generalizados são utilizados para produzir a superfície da árvore a partir do esqueleto.

É importante notar que a reconstrução de plantas a partir de nuvens esparsas de pontos tridimensionais exige alguma heurística para estimar a direção seguida pelos vários ramos da estrutura. Livny et al. (2010) definiram uma função de custo que suaviza a curvatura dos ramos e estima globalmente a melhor orientação para cada ponto da estrutura. Já Preuksakarn et al. (2010) utilizaram os pontos da nuvem como atratores para determinar as direções de crescimento do esqueleto.

## Digitalização de plantas por visão estéreo

Atualmente, equipamentos LiDAR custam dezenas de milhares de dólares. O tempo de varredura pode ser longo, caso seja desejado um grande nível de detalhe. A vantagem da visão estéreo é que alta resolução pode ser obtida com câmeras de custo inferior a 1000 dólares, facilmente encontradas no mercado. Informação de cor e textura, geralmente indisponível em sistemas LiDAR, é fornecida pelas câmeras e pode ser utilizada na modelagem e na análise dos espécimes.

---

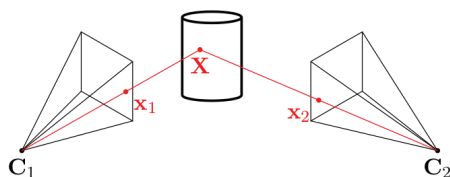
<sup>4</sup> No caso um HDS 2500 (Leica Geosystems AG, Heerbrugg, Suíça).



## Noções básicas de visão estéreo

Um dos problemas centrais da visão computacional é a recuperação da estrutura tridimensional de uma cena a partir de um conjunto de imagens bidimensionais (MA et al., 2010). A metodologia para a construção de modelos 3D pode ser generalizada em três passos essenciais (HARTLEY; ZISSERMAN, 2004).

1. Correspondência. Dado um ponto na imagem obtida por uma câmera, determinar a localização do ponto correspondente em uma segunda imagem, produzida por uma segunda câmera (ou pela mesma câmera em outra posição), como ilustrado na Figura 1.



**Figura 1.** Visão estéreo. Um objeto é visto por duas câmeras com centros de projeção  $C_1$  e  $C_2$ . O ponto  $X$ , pertencente à superfície do objeto, é visto por ambas as câmeras e é imageado na posição  $x_1$  no plano de imagem de cada câmera  $C_i$ . Se os

parâmetros de cada câmera são conhecidos (a posição do centro de projeção  $C_i$  constitui parte desses parâmetros) e se é sabido que  $x_1$  e  $x_2$  correspondem ao mesmo ponto, então a posição de  $X$  pode ser determinada por triangulação (HARTLEY; ZISSERMAN, 2004).

2. Calibração das câmeras. Dado um conjunto de pontos correspondentes, determinar a calibração das câmeras. Por calibração entendem-se os parâmetros “intrínsecos” (entre eles foco da câmera) e os parâmetros “extrínsecos” (posição e orientação da câmera).
3. Geometria da cena (estrutura). Dadas as correspondências e a calibração da câmera, é possível determinar a posição em 3D do ponto na cena que produziu cada par de correspondências pela triangulação.

Para determinar a localização 3D de um ponto, é necessário identificar sua posição em ao menos, duas imagens diferentes e obter a calibração das câmeras que produziram tais imagens. A calibração pode ser obtida por duas metodologias diferentes. A primeira delas se baseia no uso de artefatos de calibração (TSAI, 1987; ZHANG, 2000), objetos de geometria

conhecida inseridos na cena que facilitam a determinação de um conjunto de correspondências. A outra forma de obter a calibração é empregar um conjunto de correspondências determinadas, automaticamente, e técnicas robustas de estimação, como RANDOM SAMPLE CONSENSUS, RANSAC (HARTLEY; ZISSERMAN, 2004). Essas correspondências costumam ser determinadas em duas etapas: a) a detecção e representação de pontos de interesse em cada imagem e b) o uso de uma medida de similaridade na produção de um conjunto de correspondências entre esses pontos nas duas imagens (SEITZ et al., 2006). Embora dispense o uso de artefatos, essa metodologia requer um grande número de correspondências para produzir calibrações precisas, compensando por estimação robusta as correspondências espúrias produzidas pelo sistema.

### Reconstrução estéreo de plantas

O trabalho de Ivanov et al. (1995) é, possivelmente, o primeiro trabalho na literatura a utilizar visão estéreo para reconstruir a superfície tridimensional de um cultivar para fins de medição e análise. Nesse trabalho, os autores estimam a posição e orientação de folhas de milho (*Zea mays* L.) e a distribuição de sua área foliar. O sistema utilizava um par de câmeras instalado a 8.5 m do solo, sobre um milharal que apresentava altura de 2.5 m. Infelizmente, as dificuldades impostas pelo equipamento de aquisição da época (a fotografia digital ainda não havia sido largamente disseminada) prejudicaram os experimentos. A segmentação das folhas e a determinação de correspondências entre imagens foram realizadas manualmente, com o uso de ampliações fotográficas. Apesar das limitações, esse trabalho foi o precursor de sistemas mais recentes, como o de Biskup et al. (2007), que puderam contar com mais poder de computação e novos avanços na área de visão computacional.

Biskup et al. (2007) desenvolveram um sistema de visão estéreo baseado em duas câmeras digitais<sup>5</sup> para criar modelos tridimensionais para folhas de plantas de soja com o objetivo de analisar o ângulo de inclinação das folhas e seu movimento ao longo do dia. Dada a importância do

---

<sup>5</sup> Fora utilizadas duas câmeras EOS 350D Digital Rebel XT (Canon Co. Ltd, Tóquio, Japão) com lentes EF 50 mm f/1.8.

movimento para o experimento, o sistema era capaz de processar até três imagens por segundo, recuperando a informação 3D necessária para o cômputo da inclinação. Uma vez fixado o par de câmeras, a calibração foi obtida pelo algoritmo de Zhang (2000) e um artefato de calibração (no caso, um padrão xadrez impresso é fixado a um anteparo plano). O sistema estéreo produz um conjunto denso de pontos, utilizando retificação de imagens (HARTLEY; ZISSERMAN, 2004) e correlação entre vizinhanças de pixels para estabelecer o conjunto de correspondências necessário para determinar a posição em 3D de cada par de pixels pela triangulação estéreo. Para validar o método, os autores fixaram algumas folhas contra um anteparo plano e, utilizando um inclinômetro, obtiveram valores de referência em diversas inclinações. O desvio médio entre os ângulos obtidos pelo sistema estéreo proposto e os ângulos de referência foi de 1.9 +/- 0.3°.

A comunidade de visão computacional vem abordando o problema de reconstrução 3D de plantas com o objetivo de produzir modelos realistas de árvores e arbustos para computação gráfica. Porém, nesses trabalhos o objetivo é apenas obter um modelo tridimensional “verossímil”, capaz de iludir um observador quanto ao realismo da cena. As representações tridimensionais de troncos e ramos principais obtidas são fiéis ao espécime imageado. Já ramos secundários e folhas são gerados de forma procedural e apenas ajustados às imagens, de modo a reproduzir a aparência geral da planta. Embora esses métodos não produzem modelos fidedignos, algumas simulações em modelagem funcional-estrutural podem empregar modelos simplificados de plantas. Um exemplo são simulações de interceptação de luz (LOUARN et al., 2010).

Shlyakhter et al. (2001) tomaram um conjunto de fotografias de uma mesma árvore e as segmentaram manualmente, determinando quais pixels pertenciam à planta. A partir das silhuetas da árvore em cada imagem e da calibração em cada câmera, conhecida previamente, um algoritmo de intersecção volumétrica foi aplicado, determinando o “casco” do objeto (*visual hull*). Um esqueleto foi determinado a partir do casco e então utilizado como axioma (inicialização) de um sistema-L (LINDENMAYER; PRUSINKIEWICZ, 1996), que iterativamente crescia e ramificava a estrutura até preencher o volume restante no interior do casco. Apesar dos fundamentos biológicos dos sistemas-L, o método não visava a fidelidade do modelo gerado à arquitetura do espécime analisado.

Uma alternativa para calibração automática de câmeras e estimação de uma nuvem esparsa de pontos em 3D é o emprego de técnicas de estruturação por movimento (structure from motion - SfM). SfM se baseia em um resultado notável, conhecido como teorema da reconstrução projetiva (HARTLEY; ZISSERMAN, 2004), pelo qual se pode computar a reconstrução projetiva de uma cena a partir de correspondências entre imagens apenas, sem a necessidade de conhecer nada sobre a posição das câmeras envolvidas. As técnicas de SfM simultaneamente estimam a posição da câmera e a estrutura e posição dos objetos em cena.

Técnicas de SfM costumam produzir um conjunto esparsa de pontos em 3D. Tais pontos podem ser o suficiente para a reconstrução de objetos regulares como prédios e móveis que, em geral, são bem determinados pelos planos e pelas vértices (cantos). Como discutido anteriormente, objetos mais sofisticados, tais como plantas, requerem um conjunto “denso” de pontos, capazes de amostrar com certa precisão sua superfície. Quan et al. (2006) e Tan et al. (2007) utilizam como entrada de seus sistemas uma nuvem densa de pontos produzida pela técnica proposta por Lhuillier e Quan (2005). Essa técnica utiliza SFM para estimar a posição das câmeras e um conjunto esparsa inicial de pontos. Em seguida, empregando um algoritmo de crescimento de regiões (LHULLIER; QUAN, 2002), um conjunto denso de correspondências é produzido. A partir desse conjunto denso de pontos em 3D, Quan et al. (2006) combinaram técnicas de aglomerados (*clustering*), segmentação de imagens e modelos poligonais de folha para, em um processo iterativo no qual o usuário auxiliava o sistema, produzir um modelo 3D para a folhagem. O usuário utilizava, então, um editor que lhe permitia traçar ramos que completavam o modelo de plantas pequenas. Tan et al. (2007) trabalharam na direção oposta: seu sistema, baseado nos mesmos algoritmos de Lhuillier e Quan (2005), era capaz de recuperar um modelo tridimensional para troncos e ramos principais em árvores enquanto que ramos menores e folhas eram gerados artificialmente.

## Avanços recentes em reconstrução 3D de objetos

A área de reconstrução 3D por múltiplas imagens multi-view stereo (MVS) obteve grandes avanços na última década (SEITZ et al., 2006).

Recentemente, reconstrução em tempo real tem se mostrado viável em *commodity hardware* (DAVISON; NEWCOMBE, 2010; DU et al., 2011; HERNÁNDEZ; VOGIATZIS, 2011), viabilizando a construção de modelos 3D de forma interativa a partir de equipamento acessível. Uma abordagem utilizada frequentemente é recuperar a calibração das câmeras e um conjunto esparsa de pontos tridimensionais com o uso de SfM e a partir daí utilizar técnicas de crescimento de regiões para produzir uma amostragem densa para a superfície dos objetos, utilizando o conjunto esparsa como inicialização e impondo restrições baseadas em geometria projetiva, uma vez que a posição das câmeras é conhecida.

Técnicas de SfM em tempo real para uma câmera de vídeo com movimentação livre são conhecidas pelo nome de Simultaneous Localization and Mapping (SLAM) - localização e registro simultâneos. Nos últimos anos, as comunidades de robótica e realidade aumentada aprimoraram as técnicas de SLAM, produzindo sistemas capazes de determinar a posição da câmera em múltiplas escalas em tempo real de forma robusta (KLEIN; MURRAY, 2007). Newcombe e Davison (2010) calibraram uma câmera utilizando o sistema SLAM proposto por Klein e Murray (2007) para em seguida produzir uma reconstrução 3D de objetos em uma cena de escritório. Embora a posição da câmera fosse estabelecida em tempo real, alguns segundos eram necessários para que o sistema produzisse as superfícies dos objetos imageados. O método gerava uma nuvem densa de pontos em 3D pela fusão de vários mapas de profundidade produzidos entre pares de câmera Szeliski (2010). Superfícies aproximadas foram geradas a partir de um conjunto esparsa de pontos. As superfícies finais foram produzidas deformando-se a aproximação inicial com o uso do conjunto denso de pontos.

Labatut et al. (2007) utilizaram a metodologia SIFT (LOWE, 2004) para detectar pontos de interesse e determinar suas correspondências entre imagens. Restrição projetiva foi empregada para reduzir o espaço de busca ao estabelecer pares de correspondências entre pixels<sup>6</sup> que foram então utilizadas para determinar diversos pontos em 3D. O conjunto de pontos foi então empregado na construção de uma triangularização de Delaunay, que particiona o espaço 3D em tetraedros. Uma inovação do método ao

---

<sup>6</sup> Como é comum em visão estéreo, utilizou-se a restrição epipolar: dado um ponto na imagem produzida por uma câmera, a posição de seu correspondente está restrita a uma linha na imagem de. Ver Hartley; Zisserman, (2004) para detalhes.

utilizar a malha de Delaunay é particionar o espaço de forma irregular. Tal partição, se comparada a um conjunto regular de *voxels*, apresenta duas vantagens: a) maior eficiência no uso de memória, pois grandes regiões vazias no espaço são representadas por poucos tetraedros com grandes dimensões; b) um maior número de tetraedros pequenos é utilizado em regiões da cena com grande nível de detalhe.

A superfície do objeto foi obtida por um problema binário de classificação. Cada tetraedro foi classificado como interno ou externo ao objeto. A superfície foi determinada por uma malha triangular tal que cada triângulo é a face comum entre um tetraedro interno e um tetraedro externo. Tal problema de classificação foi formulado como um corte em um grafo dirigido *s-t* que minimiza uma função de energia (BOYKOV et al., 2001). Falsos positivos obtidos na geração dos pontos foram filtrados por esse processo de otimização. Três termos na função de energia foram utilizados de modo a tratar: a) oclusões; b) fotoconsistência; c) suavidade da superfície.

Vogiatzis e Hernández (2011) construíram um sistema capaz de operar em tempo real a 60 Hz. Seu método estimava a posição de uma câmera livre utilizando um artefato de calibração e estabelecia correspondências pela Normalized Cross Correlatio (NCC). Para cada pixel de cada quadro, múltiplas alternativas de correspondência foram avaliadas simultaneamente por uma elegante formulação probabilística, embora o número de alternativas fosse reduzido a uma janela de busca, o que possibilitou a operação em tempo real. Cada correspondência implicava em uma estimativa para a distância entre o ponto tridimensional  $X$ , imageado no pixel  $x_i$ , e o centro de projeção  $C_i$ . É por isso que os autores se referiam ao seu método como um “sensor de profundidade probabilístico”: a posição final de cada ponto  $X$  era definida como o máximo a posteriori de tal distribuição de profundidade.

Furukawa e Ponce (2010) empregaram a ideia de *surfels*, pequenos “retalhos” retangulares orientados no espaço 3D que são tangentes à superfície dos objetos que estão sendo reconstruídos. Projetando os *surface element* (*surfels*) nas imagens de entrada e utilizando interpolação bilinear, os autores foram capazes de avaliar a fotoconsistência entre as imagens sem assumir superfícies fronto-paralelas. Isso tornou seu método mais robusto às variações de pose dos objetos ao mesmo tempo em que mantinha o problema tratável devido à simplificação oferecida pelos *surfels*.

Um primeiro conjunto esparsa de *surfels* foi determinado pelo detector de Harris e da diferença de Gaussianas. Correspondências entre pontos fo-

ram estabelecidas empregando-se restrição epipolar. Cada par de pontos correspondentes foi utilizado para, via triangularização, inicializar um novo *surfel*. A posição e orientação do *surfel* foram refinadas, com o auxílio dos testes de fotoconsistência via NCC. Após o refinamento, o *surfel* era adicionado ao conjunto inicial se ele apresentasse boa fotoconsistência em um número suficiente de imagens.

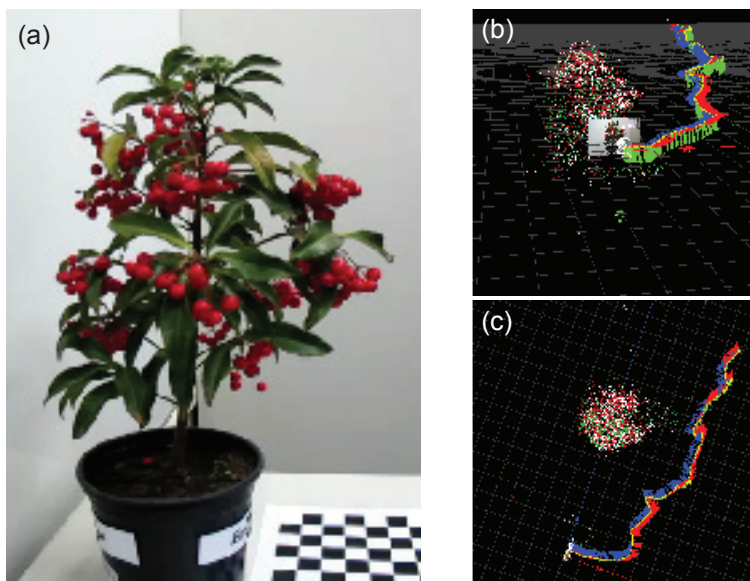
Para obter um conjunto denso de *surfels*, capaz de fornecer uma boa cobertura da superfície dos objetos, o método de Furukawa e Ponce (2010) adicionava novos *surfels* ao conjunto inicial, iterando fases de expansão e filtragem. Na fase de expansão, eram adicionados novos *surfels*, vizinhos aos já existentes, cobrindo regiões até então descobertas dos objetos imageados. Na fase de filtragem, parte dos novos *surfels* era removida por ser inconsistente: eles ocluíam ou eram oclusos por *surfels* já existentes. De acordo com os testes do comparativo de Middlebury (SEITZ et al., 2006, 2011), um dos mais empregados na área de reconstrução 3D, o método de Furukawa e Ponce (2010) vem se mostrando como um dos melhores da atualidade.

## Resultados preliminares utilizando visão estéreo

Para o presente estudo, foi realizado um experimento para avaliar se técnicas de SfM seriam capazes de obter a calibração de câmera e uma nuvem esparsa de pontos correspondentes a um espécime vegetal. Um espécime de *Ardisia crenata* “café de jardim” foi imageada com o uso de uma câmera Power Shot G11 (Canon Inc., Tóquio, Japão). A aquisição foi feita em vídeo, a 30 quadros por segundo, produzindo 600 quadros com resolução de 3648 x 2736 pixels. A câmera foi movida livremente de modo a cobrir diversos ângulos e assim reduzir problemas de oclusão.

Foi utilizado o aplicativo Voodoo (UNIVERSITY OF HANNOVER, 2010) para a estruturação por movimento. Pontos foram detectados pelo detector de Harris e as correspondências determinadas pelo rastreador de Kanade, Lucas e Tomasi (SHI; TOMASI, 1994). A cada 10 quadros, o erro de reprojeção foi minimizado: as matrizes de projeção da câmera e a posição dos pontos em 3D foram estimadas por máxima verossimilhança, um proce-

dimento conhecido como *bundle adjustment* (HARTLEY; ZISSERMAN, 2004). Os resultados podem ser vistos na Figura 2.



**Figura 2.** Estruturação por movimento. Localização da câmera durante uma sessão de aquisição de imagens para *Ardisia crenata* (“café de jardim”), determinada pela estruturação por movimento (structure from motion – SfM). (a) Primeira imagem de uma sequência de vídeo. (b) Localização espacial da câmera, destacando a posição onde foi obtida a primeira imagem. Os eixos X, Y e Z são representados pelos vetores em vermelho, verde e azul, respectivamente. (c) Vista área da trajetória da câmera durante a aquisição. É possível observar em (b) e (c) uma nuvem esparsa de pontos em 3D, pertencentes à planta e ao padrão xadrez na mesa.

Embora testes mais elaborados sejam necessários para validar o posicionamento encontrado para a câmera e para os pontos, esse resultado preliminar indica que a obtenção de uma nuvem esparsa de pontos para uma planta razoavelmente complexa como *Ardisia* é possível. Estimativas como altura da planta e largura do dossel poderiam ser obtidas a partir desse conjunto esparsa de pontos. Porém, a principal utilidade de tais resultados é servir de entrada para algoritmos mais complexos de reconstrução 3D, como o de Furukawa e Ponce (2010), capazes de gerar um conjunto denso



de pontos e, finalmente, uma malha tridimensional com um modelo para a superfície da planta. O padrão xadrez observado na Figura 2 (a) não foi utilizado na calibração da câmera (ele se mantém fora do campo de visão em boa parte do vídeo), somente os pontos detectados automaticamente foram utilizados. A independência de artefatos de calibração significa que a câmera pode ser movida livremente, de modo a capturar o máximo de informação para a resolução de situações de oclusão, comuns em plantas, nas quais elementos da estrutura ocluem outros elementos em determinados ângulos.

## Conclusões

Problemas de oclusão e a drástica mudança na aparência de folhas vistas em diferentes ângulos são problemas comuns na digitalização de plantas por visão estéreo. Porém, dentro de certos limites, avanços recentes na área de visão computacional como SfM/SLAM e reconstrução 3D por múltiplas imagens podem contornar essas dificuldades. Duas frentes se mostram particularmente promissoras: a) SfM capaz de operar em diversas escalas, atingindo o nível de detalhe necessário para a reconstrução automática de estruturas finas e b) SfM capaz de produzir respostas em tempo real, permitindo que o usuário obtenha, em “tempo de aquisição”, o volume de pontos necessário para recuperar todas as estruturas de interesse. O sistema apresentado por Du et al. (2011) é um exemplo de aquisição guiada interativa, no qual o sistema auxilia o operador na obtenção de um conjunto de imagens capaz de produzir uma reconstrução 3D satisfatória do objeto. Considerando-se esses avanços recentes, é possível progredir na reconstrução tridimensional de plantas, alcançar níveis maiores de automação e obter maior precisão nos modelos obtidos, viabilizando seu uso como instrumento para medição e análise de cultivares.

Na literatura é possível ver que pouco foi feito em relação a uma validação quantitativa dos modelos obtidos. Uma exceção é o trabalho de Preuksakarn et al. (2010) que representaram os modelos na forma de grafos de árvore em multiescala, multiscale tree graphs (MTG), como proposto por Godin et al. (1999). MTGs de referência foram gerados auto-

maticamente para árvores sintéticas e manualmente (através de técnicas de contato) para árvores reais. Os MTGs de referência e os obtidos pelos modelos foram comparados através de uma medida de distância de edição proposta por Ferraro e Godin (2000). Novos métodos para digitalização de plantas deveriam contemplar esforços similares de validação quantitativa.

## Referências

- BISKUP, B.; SCHARR, H.; SCHURR, U.; RASCHER, U. A stereo imaging system for measuring structural parameters of plant canopies. **Plant, Cell & Environment**, Oxford, v. 30, n. 10, p. 1299–1308, 2007.
- BOYKOV, Y.; VEKSLER, O.; ZABIH, R. Fast approximate energy minimization via graph cuts. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 23, n. 11, p. 1222–1239, 2001.
- CORMEN, T. H.; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. **Introduction to algorithms**. 3. ed. Cambridge: MIT Press, 2009. 1312 p.
- COSTES, E.; SINOQUET, H.; KELNER, J. J.; GODIN, C. Exploring within-tree architectural development of two apple tree cultivars over 6 years. **Annals of Botany**, London, v. 91, n. 1, p. 91–104, 2003.
- DU, H.; HENRY, P.; REN, X.; CHENG, M.; GOLDMAN, D. B.; SEITZ, S. M.; FOX, D. Interactive 3d modeling of indoor environments with a consumer depth camera. In: INTERNATIONAL CONFERENCE ON UBIQUITOUS COMPUTING, 13., 2011, Beijing. **Proceedings...** New York: ACM Computer Systemns, 2011.
- FERRARO, P.; GODIN, C. A distance measure between plant architectures. **Annals of Forest Science**, Washington, DC, v. 57, n. 5, p. 445–461, jun. 2000.
- FURUKAWA, Y.; PONCE, J. Accurate, dense, and robust multiview stereopsis. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 32, n. 8, p. 1362–1376, 2010.
- GIANNITRAPANI, R.; MURINO, V. Three-dimensional skeleton extraction by point set contraction. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 1999, Kobe. **Proceedings...** New York: IEEE, 1999. v. 1, p. 565–569.
- GODIN, C.; COSTES, E.; SINOQUET, H. A method for describing plant architecture which integrates topology and geometry. **Annals of Botany**, London, v. 84, n. 3, p. 343–357, 1999.

\_\_\_\_\_. **Plant architecture modelling**: virtual plants and complex systems. In: TURNBULL, C. (Ed.). *Plant architecture and its manipulation*. Oxford: Blackwell, 2005. p. 238–287. (Annual plant reviews, v. 17).

GRANIER, C.; AGUIRREZABAL, L.; CHENU, K.; COOKSON, S. J.; DAUZAT, M.; HAMARD, P.; THIOUX, J. J.; ROLLAND, G.; BOUCHIER-COMBAUD, S.; LEBAUDY, A.; MULLER, B.; SIMONNEAU, T.; TARDIEU, F. Phenopsis, an automated platform for reproducible phenotyping of plant responses to soil water deficit in arabidopsis thaliana permitted the identification of an accession with low sensitivity to soil water deficit. **New Phytologist**, Cambridge, v. 169, n. 3, p. 623–635, 2006.

HARTLEY, R.; ZISSERMAN, A. **Multiple view geometry in computer vision**. 2. ed. Cambridge: Cambridge University Press, 2004. 607 p.

HARTMANN, A.; CAUDERNA, T.; HOFFMANN, R.; STEIN, N.; SCHREIBER, F. HTPPheno: an image analysis pipeline for high-throughput plant phenotyping. **BMC Bioinformatics**, London, v. 12, n. 1, p. 148-156, 2011.

IVANOV, N.; BOISSARD, P.; CHAPRON, M.; ANDRIEU, B. Computer stereo plotting for 3-d reconstruction of a maize canopy. **Agricultural and Forest Meteorology**, Amsterdam, v. 75, n. 1-3, p. 85-102, 1995.

KAMINUMA, E.; HEIDA, N.; TSUMOTO, Y.; YAMAMOTO, N.; GOTO, N.; OKAMOTO, N.; KONAGAYA, A.; MATSUI, M.; TOYODA, T. Automatic quantification of morphological traits via three-dimensional measurement of Arabidopsis. **The Plant Journal**, Oxford, v. 38, n. 2, p. 358–365, 2004.

KLEIN, G.; MURRAY, D. Parallel tracking and mapping for small ar workspaces. In: IEEE AND ACM INTERNATIONAL SYMPOSIUM ON MIXED AND AUGMENTED REALITY, 6., 2007, Nara. **Proceedings...** New York: IEEE, 2007. p. 225-234.

LABATUT, P.; PONS, J.-P.; KERIVEN, R. Efficient multi-view reconstruction of large-scale scenes using interest points, delaunay triangulation and graph cuts. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION, 11., 2007, Rio de Janeiro. **Proceedings...** New York: IEEE, 2007. p. 1–8.

LANG, A. R. G. Leaf orientation of a cotton plant. **Agricultural Meteorology**, Amsterdam, v. 11, n. 1, p. 37-51, 1973.

LHUILIER, M.; QUAN, L. Match propagation for image-based modeling and rendering. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 24, n. 8, p. 1140–1146, 2002.

\_\_\_\_\_. A quasi-dense approach to surface reconstruction from uncalibrated images. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 27, n. 3, p. 418–433, 2005.

- LIVNY, Y.; YAN, F.; OLSON, M.; CHEN, B.; ZHANG, H.; EL-SANA, J. Automatic reconstruction of tree skeletal structures from point clouds. In: ACM SIGGRAPH ASIA, 10, 2010, Seoul. **Papers...** New York: ACM, 2010.
- LOPEZ, G.; FAVREAU, R. R.; SMITH, C.; COSTES, E.; PRUSINKIEWICZ, P.; DEJONG, T. M. Integrating simulation of architectural development and source-sink behaviour of peach trees by incorporating Markov chains and physiological organ function submodels into L-PEACH. **Functional Plant Biology**, Victoria, Australia, v. 35, n. 10, p. 761–771, nov. 2008.
- LOUARN, G.; COMBES, D.; SILVA, D. da; GODIN, C. Reconstruction methods to estimate light partitioning among individual plants in crowded canopies. In: INTERNATIONAL WORKSHOP ON FUNCTIONAL-STRUCTURAL PLANT MODELS, 6., 2010, Davis. **Proceedings...** Davis: University of California, 2010. p. 67–69.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. **International Journal of Computer Vision**, Amsterdam, v. 60, n. 2, p. 91–110, 2004.
- MA, Y.; SOATTO, S.; KOSECKA, J.; SASTRY, S. S. **An invitation to 3-d vision: from images to geometric models**. Amsterdam: Springer, 2010. 526 p. (Interdisciplinary applied mathematics, v. 26).
- NEWCOMBE, R. A.; DAVISON, A. J. Live dense reconstruction with a single moving camera. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2010, San Francisco. **Proceedings...** New York: IEEE, 2010. p. 1498–1505.
- PREUKSAKARN, C.; BOUDON, F.; FERRARO, P.; DURAND, J.-B.; NIKINMAA, E.; GODIN, C. Reconstructing plant architecture from 3d laser scanner data. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2010, San Francisco. **Proceedings...** New York: IEEE, 2010. p. 16–18.
- PRUSINKIEWICZ, P.; LINDENMAYER, A. **The algorithmic beauty of plants**. Amsterdam: Springer, 1996. 228 p. (The virtual laboratory).
- QUAN, L.; TAN, P.; ZENG, G.; YUAN, L.; WANG, J.; KANG, S. B. Image-based plant modeling. **ACM Transactions on Graphics**, New York, v. 25, n. 3, p. 599–604, 2006.
- RAKOCEVIC, M.; SINOQUET, H.; CHRISTOPHE, A.; VARLAET-GRANCHER, C. Assessing the geometric structure of a white clover (*Trifolium repens* L.) canopy using 3-d digitising. **Annals of Botany**, London, v. 86, n. 3, p. 519–526, 2000.
- REUZEAU, C.; FRANKARD, V.; HATZFELD, Y.; SANZ, A.; CAMP, W. van; LEJEUNE, P.; WILDE, C. de; LIEVENS, K.; WOLF, J. de; VRANKEN, E.; PEERBOLTE, R.; BROEKAERT, W. Traitmill?: a functional genomics platform for the phenotypic analysis of cereals. **Plant Genetic Resources**, Cambridge, v. 4, n. 1, p. 20–24, 2006.

- RUNIONS, A.; LANE, B.; PRUSINKIEWICZ, P. Modeling trees with a space colonization algorithm. In: EUROGRAPHICS WORKSHOP ON NATURAL PHENOMENA, 2007, Prague. [**Proceedings...**] Aire-la-Ville: Eurographics Association, 2007. p. 63-85.
- SEITZ, S.; CURLESS, B.; DIEBEL, J.; SCHARSTEIN, D.; SZELISKI, R. **Multi-view Stereo Benchmark**. 2011. Disponível em: <<http://vision.middlebury.edu/mview/>>. Acesso em: 3 mar. 2012.
- SEITZ, S. M.; CURLESS, B.; DIEBEL, J.; SCHARSTEIN, D.; SZELISKI, R. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 6., 2006, New York. **Proceedings...** Washington, DC: IEEE Computer Society, 2006. v. 1, p. 519-528.
- SHI, J.; TOMASI, C. Good features to track. In: IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 1994, Seattle. **Proceedings...** Washington, DC: IEEE Computer Society, 1994. p. 593-600.
- SHLYAKHTER, I.; ROZENOER, M.; DORSEY, J.; TELLER, S. Reconstructing 3D tree models from instrumented photographs. **IEEE Computer Graphics and Applications**, New York, v. 21, n. 3, p. 53-61, 2001.
- SINOQUET, H.; MOULIA, B.; BONHOMME, R. Estimating the three-dimensional geometry of a maize crop as an input of radiation models: comparison between three-dimensional digitizing and plant profiles. **Agricultural and Forest Meteorology**, Amsterdam, v. 55, n. 3-4, p. 233-249, 1991.
- SZELISKI, R. **Computer vision: algorithms and applications**. Amsterdam: Springer, 2010. 832 p. (Texts in computer science).
- TAN, P.; ZENG, G.; WANG, J.; KANG, S. B.; QUAN, L. Image-based tree modeling. **ACM Transactions on Graphics**, New York, v. 26, n. 3, jul. 2007.
- TSAI, R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. **IEEE Journal on Robotics and Automation**, New York. v. 3, n. 4, p. 323-344, 1987.
- TURNBULL, C. **Plant architecture and its manipulation**. Oxford: Blackwell, 2005. 322 p. (Annual plant reviews, v. 17).
- UNIVERSITY HANNOVER . Laboratorium für Informationstechnologie. **Voodoo Camera Tracker: a tool for the integration of virtual and real scenes**. Hannover, 2010. Disponível em: <<http://www.digilab.uni-hannover.de/docs/manual.html>>. Acesso em: 5 mar. 2012.
- VOGIATZIS, G.; HERNÁNDEZ, C. Video-based, real-time multi view stereo. **Image and Vision Computing**, Guilford, v. 29, n. 7, p. 434-441, 2011.

XU, H.; GOSSETT, N.; CHEN, B. Knowledge and heuristic-based modeling of laser-scanned trees. **ACM Transactions on Graphics**, New York, v. 26, n. 4, 2007.

ZHANG, Z. A flexible new technique for camera calibration. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, New York, v. 22, n. 11, p. 1330–1334, nov. 2000.



---

***Informática Agropecuária***